

Lightning Talk: CMS Use of Hadoop

Brian Bockelman

Hadoop = HDFS*

*for us, anyway

- Hadoop has many interesting part. In this talk, we're just interested in the Hadoop Distributed File System (HDFS).
- HDFS is a highly scalable distributed file system coming from the Apache Foundation (majority of developers come from Facebook and Yahoo) with an emphasis

Users

- Big external: Yahoo (25,000 nodes; largest cluster is 4,000 nodes @ 16PB), Facebook (largest cluster, 13PB)
- LHC T2: UCSD, Nebraska, Caltech, UERJ (?), Estonia (?)
- T3: UCD, UColorado, T3_ES_Oviedo (backup only), UCR (?)

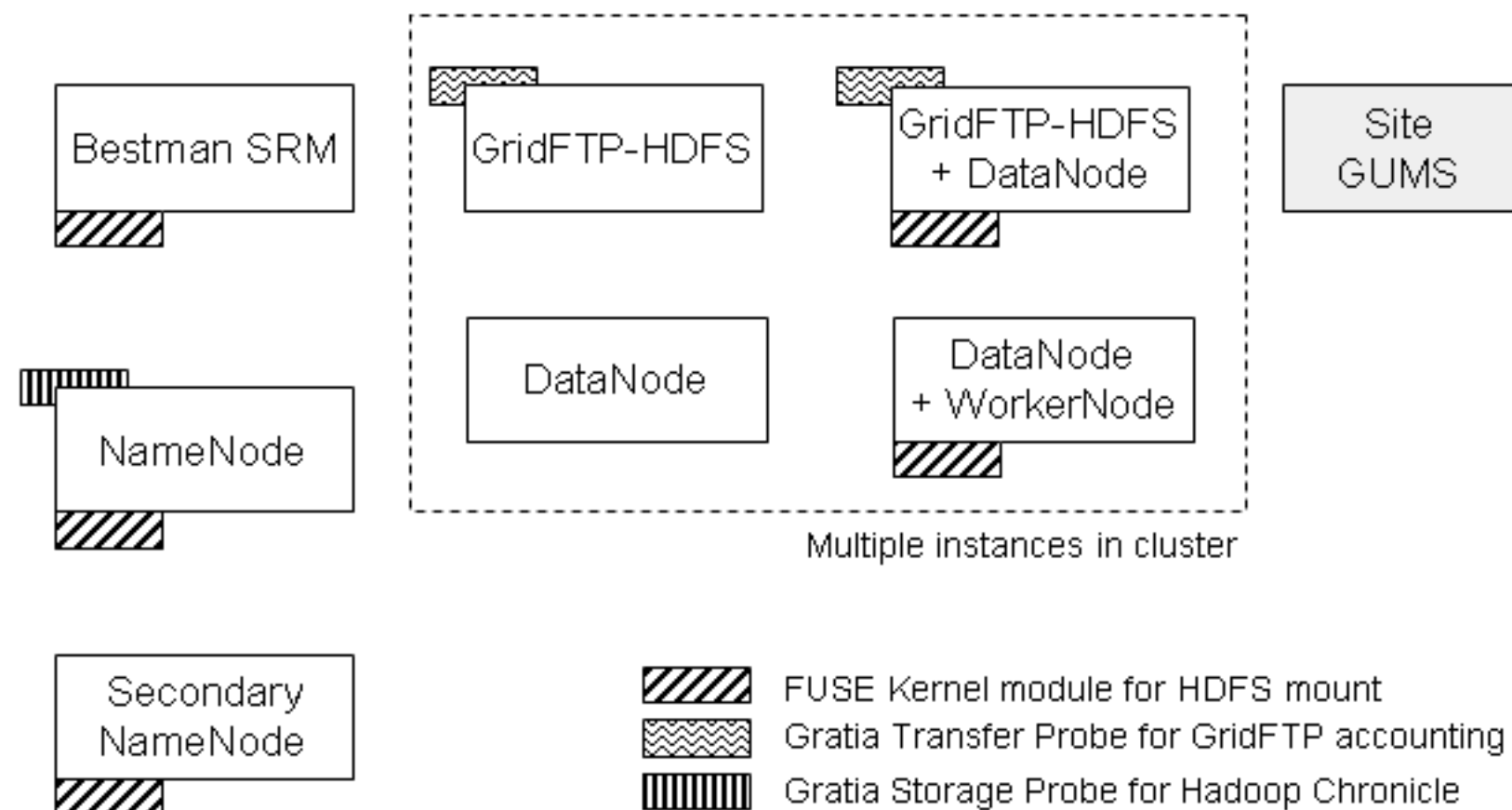
High points

- HDFS is designed to work with hard drives in worker nodes (we buy Dell r710s; 2U worker node with 6 x 2TB disks).
- Reliability is provided through replicating chunks on many datanodes.
- SRM/GridFTP provided by BestMan and Globus GridFTP, respectively.
- Completely YUM/RPM packaging is available; integrates in Linux like expected.

Focus

- The focus is on reliability:
 - Hard drive fails? No problem; data is re-replicated elsewhere.
 - Node disappears? No problem; the client re-routes its the other node holding the data.
 - Rack loses power? No problem; HDFS can keep replicas on different racks.
- We are comforted by the fact that the private-industry investment in Hadoop is \$5-10M / year.
- LHC will never be close to the largest users of this software.

Technology Details



SRM Hadoop storage system: Example topology at an OSG Site

OSG/VDT Support

- This week, we released the first official VDT version of an HDFS-based SE
 - Prior, the packaging and support was community based.
 - Support is still pretty community-based.
- The first release is expected to be supported for 6 months.
- The next release will be based on Cloudera's packaging; I expect it to last for 12 months.

Demo

- Will try a live demo now; in case it fails, youtube below:
- <http://www.youtube.com/watch?v=qoBoEzOkeDQ>